

TECHNICAL ARTICLE

A simple PCR procedure for discovering microsatellites from small insert libraries

X. W. WANG,* R. N. TRIGIANO,* M. T. WINDHAM,* R. E. DEVRIES,* B. E. SCHEFFLER,†
T. A. RINEHART‡ and J. M. SPIERS‡

*Entomology and Plant Pathology, University of Tennessee, Knoxville 37996, USA, †USDA-ARS-CGRU MSA Genomics Laboratory, 141 Experiment Station Road, Stoneville, MS 38776, USA, ‡USDA-ARS Southern Horticultural Laboratory, 810 Highway 26 West, Poplarville, MS 39470, USA

Abstract

Microsatellite discovery from genomic libraries is tedious because of the low number of clones that contain inserts and costly because of screening methodologies. A new procedure for screening clones for microsatellite DNA is described herein. Instead of colony hybridization, a polymerase chain reaction (PCR) with two vector standard primers and one synthesized repeat primer was used to directly screen colonies. PCR of colonies that produced a strong smear in gels contained the desired motif, whereas a single strong band indicated the lack of the desired motif. This simple screening method is a cost-effective way to identify microsatellite-containing colonies.

Keywords: colony hybridization, flowering dogwood, genomic library, microsatellites, PCR screening, SSR enrichment

Received 5 September 2006; revision accepted 6 November 2006

Microsatellites or simple sequence repeats (SSRs) are very useful genetic markers and are abundant and highly polymorphic in most eukaryotic nuclear genomes (Tautz 1989). Several methods for microsatellite loci isolation have been reported (Ostrand *et al.* 1992; Kandpal *et al.* 1994; Kijas *et al.* 1994; Edwards *et al.* 1996; Lench *et al.* 1996; Fischer & Bachmann 1998; Hamilton *et al.* 1999; Zane *et al.* 2002) and most involved construction of a small insert genomic library, microsatellite library enrichment, and SSR motif-containing colony selection. Based on previous studies for large-scale SSR exploitation, approximately 1–10% of the selected colonies from nonenriched genomic libraries contained SSRs, whereas about 20–75% of colonies from an enriched microsatellite library had the desired SSR motifs (Hamilton *et al.* 1999; Wang *et al.* 2004). Regardless of the way in which the libraries are prepared, colony hybridization or Southern hybridization is one of the key steps to screen SSR-containing colonies. Both techniques involve tedious work and higher costs. A more expensive alternative to hybridization is DNA-sequencing analysis of

all colonies. Recently, several (polymerase chain reaction) PCR-based methods were developed to isolate microsatellite from BAC libraries by direct sequencing using repeat-containing oligonucleotides (Fujishima-Kanaya *et al.* 2003; Waldbieser *et al.* 2003; Chen *et al.* 2005). These approaches required either a degenerate repeat-specific primer or microsatellite-flanking sequence information and resulted in a relative higher percentage of false positives. Colony PCR with vector M13 primers and a degenerated repeat-specific primer (Lunt *et al.* 1999) or nonbiotin-labelled tetranucleotide oligo (Gardner *et al.* 1999; Vadopalas & Bentzen 2000) also has been used to detect microsatellites. Those clones that produced two (or more) bands (Gardner *et al.* 1999) or a smear of products (Vadopalas & Bentzen 2000) were considered likely to contain microsatellites.

In this investigation, we used colony PCR with three primers (both forward and reverse vector primers and one synthesized complementary di- or trinucleotide repeat primer without any degenerate nucleotides as an anchor) to screen microsatellite-containing inserts. Those clones that produce a PCR smear on agarose gel are considered containing microsatellite motif. This method is rapid and efficient at screening and detecting potential microsatellites

Correspondence: R. N. Trigiano, Fax: +1-865 9744744; E-mail: rtrigian@utk.edu

from genomic libraries compared to colony or Southern hybridization and detects a high percentage of SSR-containing clones. This method may be applied to any other small insert genomic libraries.

Genomic DNA was isolated from leaves of flowering dogwood (*Cornus florida* L.), 'Cherokee Princess', using DNeasy Plant Mini Kit (QIAGEN) according to the manufacturer's instructions and then digested with several blunt-end restriction enzymes (*Alu*I, *Hae*III, or *Rsa*I) to obtain 300–800 bp fragments. The blunt-end fragments either were ligated directly into *Eco*R V-cutting pBluescript SK II (+) vector (Stratagene) (nonenriched small-insert genomic library) or ligated to adaptors, enriched with streptavidin-coated magnetic beads and then ligated into the same vector (enriched library) (Hamilton *et al.* 1999). Four SSR enriched libraries were obtained for the following motifs: (TG)₁₂ (TC)₁₂ (ATG)₈ and (AAC)₈.

White colonies growing on LB-ampicillin medium with indicator IPTG and X-Gal were placed into 96-well plates containing 200 µL LB freezing medium (Sambrook *et al.* 1989) with 100 µg ampicillin/mL and incubated at 37 °C for 15 h. Cells were stored at -80 °C or used for colony PCR amplification immediately. Ten microlitre PCR mixtures consisted of 1 µL of GeneAmp 10× PCR Buffer II (Applied Biosystems), 1 µL of 25 mM MgCl₂, 1 µL of each 2 mM of dNTPs, 0.04 µL of 5 U AmpliTaq Gold DNA polymerase (Applied Biosystems), and 3.96 µL of water, 1 µL of 2.5 mM vector forward primer T₃ (5'-AATTAAACCCTCACT-AAAGGG-3'), 1 mL of 2.5 mM vector reverse primer T₇ (5'-GTAATACGACTCACTATAGGGC-3'), 1 µL of 2.5 mM of one of the following SSR motif oligos: (TG)₁₂ for (TG)_n library (TC)₁₂ for (TC)_n library (ATG)₈ for (ATG)_n library or (AAC)₈ for (AAC)_n library. Cells were transferred with a pipette tip to the PCR master mix in 96-well plates. The PCR program was 96 °C for 5 min, followed by 35 cycles of 95 °C for 1 min, 50 °C for 1 min, 72 °C for 1 min, and 72 °C for 1 min and performed on Eppendorf Autorisierter Thermocycler. PCR products were separated on 2% agarose gels and visualized with ethidium bromide.

Plasmid DNA of 768 colonies that showed smears in agarose gels were isolated using a modified alkaline lysis method and sequenced (ABI Big-Dye version 3.1 terminators)

on a Model ABI 3730XL capillary electrophoresis DNA sequencer with 50 cm arrays.

We constructed both a small insert genomic library (nonenriched) and four repeat-enriched libraries. Primers T₃ and T₇ can amplify any fragment within the molecular cloning site (MCS) of cloning vector (Fig. 1A); therefore, when a nonmotif insert is present only one strong, distinctive band was present in the gel (Fig. 1B; lane 14). However, if one of the SSR motifs was present, the third primer annealed to different sites within the repeat area (Fig. 1A) and the various products were detected as a smear in the gel (Fig. 1B; lanes 2–13). Initially, 10 colonies from each enriched library in which PCR had revealed a smear of products were selected for sequencing and all contained the desired motif. Additionally, 10 colonies that produced only a single clear band were sequenced and none contained any repeats.

The four different enriched libraries were screened using this PCR method. An average of 46.8% (with a range of 29–64%) of the selected colonies from four enriched libraries showed smear patterns indicating the possible presence of microsatellites (Table 1). Although the rate of SSR enrichment is lower than in other studies (Wang *et al.* 2004), it is similar to others that used the method Hamilton *et al.* (1999) for enrichment. Colonies from the small insert library before enrichment were examined and approximately 2% showed smear patterns (data not shown). Thus, the Streptavidin-based enrichment methods increased the presence of microsatellites 23.4-fold on average (15–32-fold) over the original small-insert library.

From the 1034 smear-positive colonies, 768 inserts (equal to eight 96-well plates) were sequenced and 760 contained repeats; eight contained very short (< 4) repeats. Thus, about 99% of smear-positive colonies contained SSRs, which is higher than other PCR methods (Lunt *et al.* 1999; Fujishima-Kanaya *et al.* 2003; Waldbieser *et al.* 2003; Chen *et al.* 2005). In contrast, Vadopalas & Bentzen (2000) detected clones that potentially contained SSRs by smears of products in agarose gels using either the M13 forward or reverse primer with a (GATA)₇ second primer, a third primer was not used. However, only 76% (62 of 82) of the smear-positive clones contained the proper microsatellite motif. In our study, there is no significant relationship between

Table 1 Analysis of colonies containing microsatellite motifs in enriched libraries

Libraries	Motifs			
	(AC) _n	(TC) _n	(ATG) _n	(AAC) _n
No. of white colonies*	480	768	480	480
No./% of colonies with PCR product smear in agarose gel	263/54.8	490/63.8	138/28.8	143/29.8
No. colonies selected for sequencing of insert	192	384	96	96
No./% of colonies containing motif†	191/99.5	383/99.7	92/95.8	94/97.9

*White colonies on LB-ampicillin medium (Sambrook *et al.* 1989); †inserts with repeat length less than four were not included.

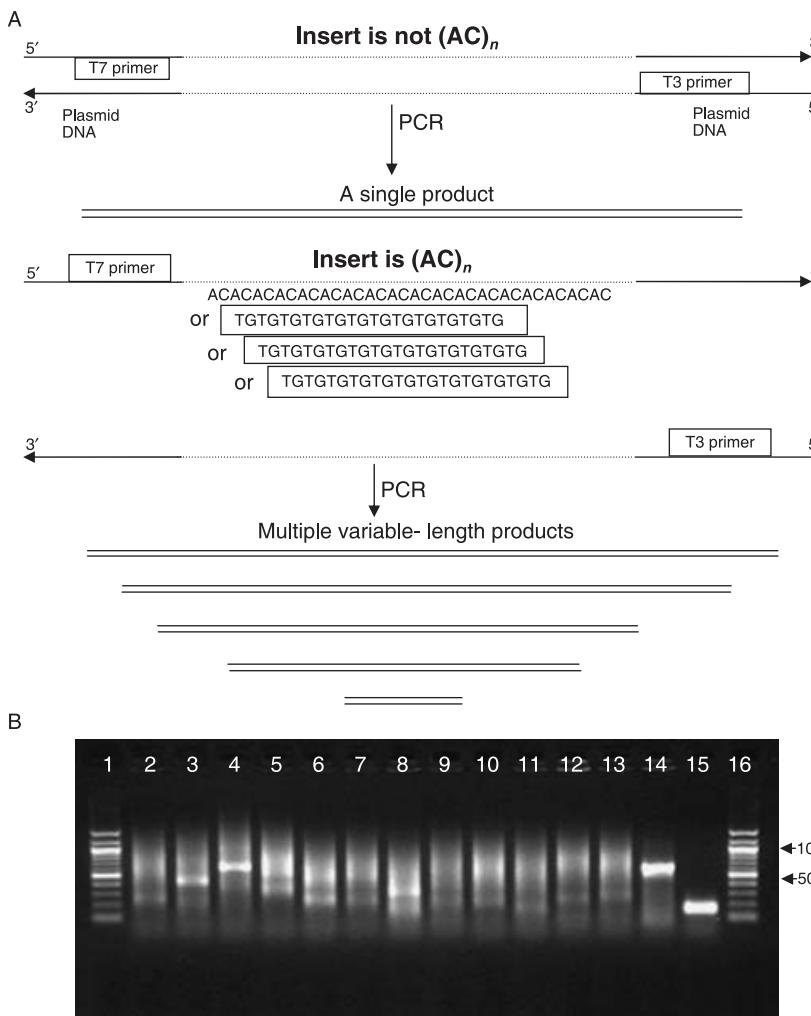


Fig. 1 Sketch map (A) of PCR amplification for detection of DNA inserts. If insert does not contain desired motif (AC), only one amplification product (sequence between two internal plasmid primers (T_3 and T_7) is possible and a single band appears in the lane (see lane 14 in B). However, if insert does contain proper motif (AC), the third primer (TG_{12}) anneals at various positions within the insert and yields variable length amplified products. The products produce a 'smear' in an agarose gel, which indicates the desired motif is present (see lanes 2–13 in B). Colony PCR amplification (B) with vector standard primers T_3 and T_7 and repeat (TG_{12}) primer from (AC)_n-enriched microsatellite library of *Cornus florida*. Lanes 2–13 indicated PCR pattern of colonies that contain microsatellite motifs – note the diffuse staining 'smears'. Lane 14 showed only one prominent band indicating an insert, but it did not contain the proper microsatellite motif. Lane 15 is the empty (without insert DNA) vector. Size standard (lanes 1 and 16) is 1 kb DNA ladder.

smear patterns and the length of repeats. However, if the smear was short and intensely stained, and was at the expected insert size (400–800 bp), the possibility of longer microsatellite with flanking sequence was increased. In contrast, if the smear was dispersed over a wider range of products and not stained strongly, it is likely that a short repeat (< 10 repeat units) was present. However, colonies producing weakly stained PCR smears should be sequenced as all contained an SSR with the proper motif (data not shown).

Traditionally, colony or Southern hybridization is used to screen microsatellites from various enriched libraries (Ostrander *et al.* 1992; Kandpal *et al.* 1994; Kijas *et al.* 1994; Edwards *et al.* 1996; Lench *et al.* 1996; Connell *et al.* 1998; Fischer & Bachmann 1998; Hamilton *et al.* 1999). Both colony and Southern hybridization require 3 days to screen positives. However, our PCR screen needs only 3 hours (excluding subculturing of clones) and is less expensive (avoids nylon membranes, labelling probes, and isotopes). Although several PCR-based methods obviated colony

hybridization, they still needed a degenerate repeat-specific primer and/or sequence information to design the reverse primer. These techniques also generated a relatively higher proportion of false positives (Lunt *et al.* 1999). Nondegenerate primers also have been used in the PCR colony screening either producing two or more bands (Gardner *et al.* 1999) or a smear of products (Vadopalas & Bentzen 2000). However, using a two-primer system, 25% false positive clones with a tetranucleotide insert motif were identified (Vadopalas & Bentzen 2000). The present system is almost 100% correct for identifying SSR motif-containing colonies. In contrast, isolation of microsatellites from genomic libraries without enrichment by direct sequencing using repeat-containing oligonucleotides is easy, but identifies a very lower percentage of insert positive colonies (Connell *et al.* 1998; Adcock & Mulder 2002).

With the present study, 99% of the colonies that produced a short, strong PCR smear contained microsatellite repeat sequences. The detection of higher percentages of positive

colonies decreases the cost of subsequent sequencing for most small laboratories. Furthermore, PCR screening is cost-effective compared to randomly sequencing colonies from microsatellite-enhanced libraries, especially given the variable success rates (25–65%) for enriched libraries. For example, if the cost of bidirectional sequencing is \$5 per insert and the cost of our PCR screen is \$0.20 per reaction, then identifying 1000 colonies from an enhanced library containing 50% microsatellites would cost \$5000 using DNA sequencing alone and only \$2700 using the PCR screen [(1000 colonies × \$0.20 PCR) + (\$2500 colonies bi-directional sequencing) vs. 1000 colonies × \$5 for bi-directional sequencing]. Cost savings increase when percentage of colonies containing desired inserts goes down, which makes our PCR screen better suited for genomes where library construction or SSR enrichment is difficult.

This simple PCR screen for SSR-containing colonies has the following two main advantages for small research groups: (i) there is no need for colony hybridization to detect desired inserts; and (ii) substantial time and cost savings are realized by not having to sequence inserts that do not contain the proper motif. The colony PCRs are 99% effective in detecting SSRs and are easy to complete in minimally equipped laboratories.

Acknowledgements

This work was supported by USDA Agreement no. 58-6404-2-0057. We thank Mary Duke, Linda Ballard and Xiaofen Liu for technical assistance.

References

- Adcock GJ, Mulder RA (2002) Polymorphic microsatellite loci for paternity analysis in the Madagascar paradise flycatcher (*Terpsiphone mutata*: Aves). *Molecular Ecology Notes*, **2**, 287–289.
- Chen K, Knorr C, Bornemann-Kolatzki K *et al.* (2005) Targeted oligonucleotide-mediated microsatellite identification (TOMMI) from large-insert library clones. *BMC Genetics*, **6**, 54. <http://www.biomedcentral.com/1471-2156/6/54>.
- Connell JP, Pammi S, Iqbal MJ *et al.* (1998) A high throughput procedure for capturing microsatellites from complex plant genomes. *Plant Molecular Biology Report*, **16**, 341–349.
- Edwards KJ, Barker JH, Daly A *et al.* (1996) Microsatellite libraries enriched for several microsatellite sequences in plants. *BioTechniques*, **20**, 758–760.
- Fischer D, Bachmann K (1998) Microsatellite enrichment in organisms with large genomes (*Allium cepa* L.). *BioTechniques*, **24**, 796–802.
- Fujishima-Kanaya N, Toki D, Suzuki K *et al.* (2003) Development of 50 gene-associated microsatellite markers using BAC clones and the construction of a linkage map of swine chromosome 4. *Animal Genetics*, **34**, 135–141.
- Gardner MG, Cooper SJB, Bull CM, Grant WN (1999) Isolation of microsatellite loci from a social lizard, *Egernia stokesii*, using a modified enrichment procedure. *Journal of Heredity*, **90**, 301–304.
- Hamilton MB, Pincus EL, Di Fiore A *et al.* (1999) Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *BioTechniques*, **27**, 500–507.
- Kandpal RP, Kandpal G, Weissman SM (1994) Construction of libraries enriched for sequence repeats and jumping clones, and hybridization selection for region-specific markers. *Proceeding National Academy of Sciences, USA*, **91**, 88–92.
- Kijas JM, Fowler JC, Garbett CA *et al.* (1994) Enrichment of microsatellites from the citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. *BioTechniques*, **16** (656–660), 662.
- Lench NJ, Norris A, Bailey A *et al.* (1996) Vectorette PCR isolation of microsatellite repeat sequences using anchored dinucleotide primers. *Nucleic Acids Research*, **24**, 2190–2191.
- Lunt DH, Hutchinson WF, Carvalho GR (1999) An efficient method for PCR-based isolation of microsatellite arrays (PIMA). *Molecular Ecology*, **8**, 891–894.
- Ostrander EA, Jong PM, Rine J *et al.* (1992) Construction of small-insert genomic DNA libraries highly enriched for microsatellite repeat sequences. *Proceeding National Academy of Sciences, USA*, **89**, 3419–3434.
- Sambrook J, Fritsch EF, Maniatis T (1989) *Molecular cloning: a Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Tautz D (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*, **17**, 6463–6471.
- Vadopalas B, Bentzen P (2000) Isolation and characterization of di- and tetranucleotide microsatellite loci in geoduck clams, *Panopea abrupta*. *Molecular Ecology*, **9**, 1435–1436.
- Waldbieser GC, Quiniou SM, Karsi A (2003) Rapid development of gene-tagged microsatellite markers from bacterial artificial chromosome clones using anchored TAA repeat primers. *BioTechniques*, **35**, 976–979.
- Wang XW, Kaga A, Tomooka N *et al.* (2004) The development of SSR markers by a new method in plants and their application to gene flow studies in azuki bean [*Vigna angularis* (Willd.) Ohwi & Ohashi]. *Theoretical Applied Genetics*, **109**, 352–360.
- Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology*, **11**, 1–16.